PCT/EP99/06302

Europäisches Patentamt | European Patent Office | Office européen des brevets



# Bescheinigung | Certificate | Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

**Patentanmeldung Nr. | Patent application No. | Demande de brevet n°**

98202854.0

## PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Der Präsident des Europäischen Patentamts:
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

I.L.C. HATTEN-HECKMAN

DEN HAAG, DEN
THE HAGUE, 11/10/99
LA HAYE, LE

Eur päisches Patentamt
European Patent Office
Office européen des brevets

# Blatt 2 der Bescheinigung
# Sheet 2 of the certificate
# Page 2 de l'attestation

Anmeldung Nr.:
Application no.: 98202854.0
Demande n°:

Anmeldetag:
Date of filing: 27/08/98
Date de dépôt:

Anmelder:
Applicant(s):
Demandeur(s):
Wallaart, Thorvald Eelco
9727 DJ Groningen
NETHERLANDS

Bezeichnung der Erfindung:
Title of the invention:
Titre de l'invention:
DNA encoding amorpha-4,11-diene synthase

In Anspruch genommene Priorität(en) / Priority(ies) claimed / Priorité(s) revendiquée(s)

| Staat: State: Pays: | Tag: Date: Date: | Aktenzeichen: File no. Numéro de dépôt: |
|---|---|---|

Internationale Patentklassifikation:
International Patent classification:
Classification internationale des brevets:
C12N15/60, C12N15/70, C12N15/82, C12N9/88, C12N5/10, C12N1/21, C12P5/00, C12P17/18, A01H5/00

Am Anmeldetag benannte Vertragstaaten:
Contracting states designated at date of filing: AT/BE/CH/CY/DE/DK/ES/FI/FR/GB/GR/IE/IT/LI/LU/MC/NL/PT/SE
Etats contractants désignés lors du depôt:

Bemerkungen:
Remarks:
Remarques:

EPO - DG 1

27. 08. 1998

# AMORPHA-4,11-DIENE

The present invention relates to a DNA sequence, a polypeptide encoded by this sequence, and to the use of said DNA sequence and polypeptide.

Human malaria is a commonly occurring widespread infectious disease, caused in 85 % of the cases by Plasmodium falciparum. This parasite is responsible for the most lethal form of malaria, malaria tropicana. Each year, malaria causes clinical illness, often very severe, in over 100 million people of which eventually over 1 million individuals will die. Approximately 40% of the world's population is at risk of malaria infection (as estimated by the World Health Organization).

Malaria has traditionally been treated with quinolines, such as quinine, chloroquine, mefloquine and primaquine, and with antifolates. Unfortunately, most P.falciparum strains have become resistant to chloroquine, and some have developed resistance to mefloquine and halofantrine as well. Thus, novel antimalarial drugs to which resistant parasites are sensitive are urgently needed. Artemisinin, as well as its semisynthetic derivatives are promising candidates here.

Artemisinin (Fig. 1), [3R-(3α,5aß,6ß,8aß, 9α,12ß,12aR*)]-Octahydro-3,6,9-trimethyl-3,12-epoxy-12H-pyrano[4,3-j]-1,2-benzodioxepin-10(3H)-one; molecular weight 282.35), also called arteannuin, qinghaosu or QHS, is a sesquiterpene lactone endoperoxide isolated from the aerial parts of the plant Artemisia annua L.

Artemisia annua L., also known as quinghao (Chinese), annual or sweet wormwood, or sweet annie is an annual herb native to Asia. A.annua, a member of the Asteraceae, belongs to the tribe Anthemideae of the Asteroideae, and is a large herb often reaching more than 2.0 m in height. It is usually single-stemmed with alternating branches. The aromatic leaves are deeply dissected and range from 2.5 to 5 cm in length. Artemisinin is mainly produced in the leaves as a

secondary metabolite at a concentration of 0.01 - 0.6 % on a dry weight base in natural populations. Artemisinin is unique to the plant A.annua with one possible exception of A.apiacea L. The A.annua used in this invention is of Vietnamese origin.

Because of its low concentration in plants, artemisinin is a relatively expensive resource for a drug. Current research has thus been aimed at producing artemisinin at a larger scale by organic synthesis. However, because artemisinin consist of seven chiral carbon atoms, theoretically $2^7 = 128$ isomers can be formed of which only one is identical to artemisinin. Because of this complex structure of artemisinin, production of this compound by organic synthesis is not profitable from a commercial point of view.

Genetic engineering of the biosynthetic pathway of artemisinin may give rise to higher artemisinin levels in plants. To be able to interfere in the biosynthesis of artemisinin, the biosynthetic pathway has to be known, either completely or partially. Several attempts to elucidate the entire biosynthetic pathway have been undertaken. Until now, however, the exact pathway has remained largely unknown.

In the research that led to the present invention, a unique pathway has been discovered which has not been published before. Thus, the formation of amorpha-4,11-diene (1ß,6ß,7ß,10αH-amorpha-4,11-diene) and the hydroperoxide of dihydroarteannuic acid in A.annua has not been reported before in literature.

From literature it is known that terpene cyclases (synthases) are branch point enzymes, which likely play an important role in terpenoid biosynthesis. Overexpression of such a branch point enzyme (terpene cyclase) may increase terpenoid production in an organism. Factors that may influence the success of such an approach are, in the case of artemisinin, the number and nature of the subsequent biosynthetic steps leading to artemisinin. Figure 2 shows the postulated

biosynthetic pathway of artemisinin (based on literature and personal data). This pathway is divided into three parts:

The first part (Part I) represents the terpenoid (Isoprenoid) pathway. This pathway is a general pathway. Farnesyl diphosphate (farnesyl pyrophosphate) (FPP), for example, is present in every living organism and it is the precursor of a large number of primary and secondary metabolites. It has been established that FPP is the precursor of all sesquiterpenes. Thus, by definition FPP is the precursor of artemisinin.

Part II displays the cyclization of the general precursor FPP into the highly specific precursor amorpha-4,11-diene (further referred to as amorphadiene), the first specific precursor of artemisinin. In this pathway amorphadiene synthase is a branch point enzyme, having a key position in the biosynthetic pathway of artemisinin.

In part III, dihydroarteannuic acid (DHAA), also called dihydroartemisinic acid, is photo-oxidatively converted into its hydroperoxide (DHAA-OOH). This hydroperoxide of DHAA will spontaneously oxidize into artemisinin. No enzymes are involved in this part of the pathway and therefore it is impossible to alter artemisinin production by overexpression of genes involved in this part of the pathway.

Cytochrome P-450 catalyzed enzymes and an enoate reductase are probably involved in the conversion of amorphadiene into DHAA, the transition state between part II and part III (see figure 2[A]). Because no intermediates of this part of the pathway are known or present (accumulated) in detectable amounts, in the plant, (except arteannuic acid, also called artemisinic acid or 4,11(13)-amorphadien-12-oic acid) it is likely that these precursors are very rapidly converted into DHAA. A rate limiting step in this part of the pathway is not very likely.

Taking al these aspects into account the most logical step to be altered by genetic interfering, is the

conversion (cyclization) of FPP into amorphadiene by amorphadiene synthase.

The object of the present invention is therefore to provide a DNA sequence which exhibits at least a 70% homology to the sequence as shown in figure 9a, and which codes for a polypeptide having the biological activity of the enzyme amorphadiene synthase.

The biological activity of the enzyme amorphadiene synthase relates to the conversion of the general precursor farnesyl pyrophosphate (FPP) into the specific precursor amorphadiene, which is further converted to artemisinin in A.annua.

By transforming a suitable host cell with the DNA sequence of the invention, the conversion of farnesyl pyrophosphate (FPP) into the highly specific precursor amorphadiene can be increased. Suitable host cells are for example bacterial cells, such as E.coli, or plant cells such as those of A.annua.

By adding FPP to a culture medium further comprising the polypeptide of the invention (as described in example I), or transformed cells, e.g. E.coli, comprising the DNA sequence of the invention (as described in examples III and IV), FPP is converted into amorphadiene.

In addition, several plants are capable of producing large amounts of FPP making them potential organisms for amorphadiene production.

Transformed cells can be used either in disrupted form, by for example sonication, or as intact cells, as a source of amorphadiene.

Overexpression of the amorphadiene synthase encoding gene in A.annua will increase artemisinin production, because the terpene cyclase is expected to be the rate limiting step.

The results of the present research (postulated biosynthetic pathway of artemisinin) make the presence of a single major rate limiting step at the place of the amorphadiene synthase very likely. Overexpression of the

amorphadiene synthase encoding gene can increase the production of artemisinin in A.annua.

The chemical structure of the first specific precursor of artemisinin, a cyclization product of FPP, was not known in literature. Neither has anyone so far detected such a compound in A.annua. Nevertheless it was possible to predict a likely structure for this cyclization product, based on the structure of DHAA and arteannuic acid (Fig.2[A]). The structure predicted in this way was consistent with a compound which is known in literature as 4,11-Amorphadiene (1ß,6ß,7ß,10αH-Cadina-4,11-diene, incorrect name), as depicted in figure 3. This compound, isolated from Viguiera oblongifolia, has previously been described by Bohlmann et al. (Phytochemistry, vol. **23**, 5, 1183-1184 (1984)). Starting from arteannuic acid (isolated from A.annua), it was possible to synthesize amorphadiene. Amorphadiene obtained in this way was in all chemical and physical aspects identical to amorphadiene as described by Bohlmann et al., and this standard was used to show the presence of amorphadiene in a terpene extract of A.annua.

A further object of the present invention is to provide a polypeptide having the biological activity of the enzyme amorphadiene synthase, obtainable by a process as described in example I. This polypeptide can be used to convert FPP into amorphadiene which subsequently can be converted into artemisinin in A.annua.

Amorphadiene, produced by a suitable host organism transformed with the DNA sequence of the invention as precursor, can subsequently be chemically converted to dihydroarteannuic acid. Dihydroarteannuic acid can be used as an antioxidant.

The chemical conversion of amorphadiene into dihydroarteannuic acid (Fig. 12) starts with the enantio-, stereo- and regioselective (anti-markownikoff) hydroboration of amorphadiene with $BH_3$, yielding a trialkylborane, followed by an oxidation of the trialkylborane with NaOH / $H_2O_2$ yielding the alcohol

(Advanced Organic Chemistry, Jerry March, 4th Edition, Wiley, 1992). A mild oxidation of the alcohol to the acid can be obtained by PDC (pyridinium dichromate) without attacking the second double bond (Fig. 12) (Organic Synthesis, M.B. Smith, 1st Edition, McGraw-Hill, 1994).

Many genes encoding enzymes involved in the biosynthetic pathway of farnesyl diphosphate are cloned and known in literature. For A.annua, for example, the sequence of the farnesyl diphosphate synthase encoding gene is known in literature (Y. Matsushita, W-K. Kang and V. Charlwood Gene, 172 (1996) 207-209). A further approach to introduce or increase the amorphadiene production in an organism, is to transform such an organism (for example A.annua) simultaneously with the DNA sequence of the invention with one or more genes involved in the biosynthesis of farnesyl diphosphate. The expression of a fusion protein of amorphadiene synthase and farnesyl diphosphate synthase may be an example here.

(Sesqui)terpenes, such as amorphadiene, are also known as flavor and fragrance compounds in the food and perfume industry. In addition, terpenes play a role in plant-insect interactions, such as the attraction or repulsion of insects by plants. In addition, dihydroarteannuic acid, which is an intermediate in the metabolic route from amorphadiene into artemisinin in A.annua, can be used as an antioxidant.

Amorphadiene, obtained by (over)expression of the DNA sequence of the invention, or by using the polypeptide of the invention, can be applied for these purposes as well.

The plants that can be used for this invention are preferably plants already producing artemisinin. A prime example is Artemisia annua, as this species contains the remainder of the pathway leading to artemisinin. However, this invention may also be used for the production of amorphadiene in plants, which, as mentioned before, can be used as a flavor or fragrance compound or biocide, or can be converted to artemisinin,

either chemically or by bioconversion using microorganisms or plant cells.

The plant that can be used for the production of amorphadiene is preferably a plant already producing sesquiterpenes, as these plants already have the basic pathway and storage compartments available, or a plant in which the biosynthesis of sesquiterpenoids can be induced by elicitation. The methods of this invention are readily applicable via conventional techniques to numerous plant species, including for example species from the genera Carum, Cichorium, Daucus, Juniperus, Chamomilla, Lactuca, Pogostemon and Vetiveria, and species of the inducible (by elicitation) sesquiterpenoid phytoalexin producing genera Capsicum, Gossypium, Lycopersicon, Nicotiana, Phleum, Solanum and Ulmus. However, also common agricultural crops like soybean, sunflower and rapeseed may be interesting candidates here.

The invention will be further illustrated by the following examples and figures, but will not be limited thereto.

**EXAMPLES**

**EXAMPLE 1**

Conversion of farnesyl pyrophosphate into amorphadiene by amorphadiene synthase

A. Isolation, partial purification and identification of amorphadiene synthase from A.annua

During enzyme isolation and preparation of the assays, all operations were carried out on ice or at 4°C. Ten grams of frozen young leaves from greenhouse-grown A.annua were ground in a pre-chilled mortar and pestle in 40 ml of pre-chilled buffer containing 25 mM MES (Ph 5.5), 20% (v/v) glycerol, 25 mM sodium ascorbate, 25 mM $NaHSO_3$, 10 mM $MgCl_2$ and 5 mM DTT (buffer A) slurried with 1 g polyvinylpolypyrrolidone (PVPP) and a spatula tip of purified sea sand. Ten grams of polystyrene resin (Amberlite XAD-4, Serva) were added and the slurry was

stirred carefully for 10 min and then filtered through cheesecloth. The filtrate was centrifuged at 20,000g for 20 min (pellet discarded), and then at 100,000g for 90 min. A 3-ml subsample of the supernatant was desalted to a buffer containing 15 mM MOPSO (pH 7.0), 10% (v/v) glycerol, 1 mM sodium ascorbate, 10 mM $MgCl_2$ and 2 mM DTT (buffer B) and used for enzyme assays/product identification (see below at 'B').

The remainder of the supernatant was added to 12.5 g DEAE anion exchanger (Whatman DE-52), which had been rinsed several times with buffer A, and stirred carefully for 10 min. After centrifugation at 18,000g for 20 min, the supernatant was decanted and the DE-52 pellet discarded. Proteins in the supernatant were precipitated by adding $(NH_4)_2SO_4$ to a final concentration of 70%, careful stirring for 30 min, and centrifugation at 20,000g for 10 min. The resulting pellet was resuspended in 6 ml buffer A and desalted to buffer B. After addition of glycerol up to 30% (v/v) this enzyme preparation could be frozen in liq $N_2$ and stored at -80°C without loss of activity. 0.5 ml of this enzyme preparation was applied to a Mono-Q FPLC column (HR5/5, Pharmacia Biotech), previously equilibrated with buffer B without sodium ascorbate, with 0.1% Tween-20. The enzyme was eluted with a gradient of 0-2.0 M KCl in the same buffer. For determination of enzyme activities, 50 $\mu$L of the 0.75-ml fractions were diluted 2-fold in an Eppendorf tube with buffer B and 20 $\mu$M [$^3$H]FPP was added. The reaction mixture was overlaid with 1 mL of hexane to trap volatile products and the contents mixed. After incubation for 30 min at 30°C, the vials were vigorously mixed, and centrifuged briefly to separate phases. A portion of the hexane phase (750 $\mu$L) was transferred to a new Eppendorf tube containing 40 mg of silica gel (0.035-0.07 mm, pore diameter 6 nm, Janssen Chimica) to bind terpenols produced by phosphohydrolases, and, after mixing and centrifugation, 500 $\mu$L of the hexane layer was removed for liquid scintillation counting in 4.5 ml of Ultima

Gold cocktail (Packard). The active fractions were combined, and an assay carried out to determine product identity (see below). After the Mono-Q step, the enzyme was separated from all other FPP-converting activities (Fig. 4C). This enzyme preparation was used for the measurement of enzyme characteristics such as molecular weight and $K_m$. The molecular weight was determined using size-exclusion chromatography. 200 µL of the Mono-Q eluent was loaded on a Superdex 75 (H/R10/30, Pharmacia Biotech) and eluted in the same buffer as used for Mono-Q. Enzyme activities in 0.5 ml fractions were determined as described for Mono-Q, but using undiluted eluens. The column was calibrated using cytochrome C, ribonuclease A, α-chymotrypsinogen, ovalbumin and BSA (all from Sigma). The estimated molecular weight was 56 kDa (Figure 13). Enzyme-kinetics were determined using 5- and 10-fold diluted Mono-Q eluted enzyme preparation and [$^3$H]-FPP concentrations ranging from 0.25-100µM. $K_m$ for amorphadiene synthase was 0.6 µM.

B. Determination of product identity

For determination of product identity, 20 µM [$^3$H]-FPP (Amersham; for radio-GC analysis) or 50 µM unlabelled FPP (Sigma; for GC-MS analysis) were added to 1 ml of the enzyme preparations. After the addition of a 1mL redistilled pentane overlay to trap volatile products, the tubes were carefully mixed and incubated for 1 h at 30°C. Boiled samples were used as controls. Following the assay, the tubes were vigorously mixed. The organic layer was removed and passed over a short column of aluminum oxide overlaid with anhydrous $MgSO_4$. The assay was extracted with another 1 mL of diethyl ether which was also passed over the aluminum oxide column, and the column washed with 1.5 mL of diethyl-ether. For GC-analysis, the combined pentane/diethyl-ether mixture was slowly concentrated under a stream of $N_2$.

Radio-GLC was performed on a Carlo-Erba 4160 Series gas chromatograph equipped with a RAGA-90

radioactivity detector (Raytest, Straubenhardt, Germany). Sample components eluting from the column were quantitatively reduced before radioactivity measurement by passage through a conversion reactor filled with platinum chips at 800°C. Samples of 1 µL were injected in the cold on-column mode. The column was a fused silica capillary (30 m x 0.32 mm i.d.) coated with a film of 0.25 µm of polyethylene glycol (EconoCap EC-WAX, Alltech Associates) and operated with a He-flow of 1.2 mL $min^{-1}$. The oven temperature was programmed to 70°C for 5 min, followed by a ramp of 5° $min^{-1}$ to 210°C and a final time of 5 min. To determine retention times and peak identities (by co-elution of radioactivity with reference standards), about 20% of the column effluent was split with an adjustable splitter to an FID (temperature 270°C). The remainder was directed to the conversion reactor and radio detector. $H_2$ was added prior to the reactor at 3 mL $min^{-1}$, and $CH_4$ as a quench gas prior to the radioactivity detector (5 mL counting tube) to give a total flow of 36 mL $min^{-1}$. The major [$^3$H]-labelled product co-eluted with the amorphadiene reference standard (ret. time 14 min) (Fig. 4B). The second radiolabelled product is farnesol, the product of aspecific phosphohydrolase activity. After the Mono-Q step, the enzyme was separated from all other FPP-converting activities (Fig. 4C). This enzyme preparation was used for the measurement of enzyme characteristics such as molecular weight and $K_m$.

GC-MS analysis was performed using a HP 5890 series II GC and HP 5972A Mass Selective Detector (Hewlett-Packard) equipped with an HP-5MS or HP-Innowax column (both 30 m x 0.25 mm i.d., 0.25 µm df). The oven was programmed at an initial temperature of 70°C for 1 min, with a ramp of 5°C $min^{-1}$ to 210°C and final time of 5 min. The injection port (splitless mode), interface and MS source temperatures were 175, 290 and 180°C, respectively, and the He inlet pressure was controlled by electronic pressure control to achieve a constant column flow of 1.0 ml $min^{-1}$. Ionization potential was set at 70

eV, and scanning was performed from 30-250 amu. The $(NH_4)_2SO_4$ precipitated enzyme preparation was free of endogenous sesquiterpenes. GC-MS analysis on the two different GC-columns of sesquiterpene products generated from FPP by this enzyme preparation showed that the main product had a mass spectrum and retention time equal to that of the semi-synthetically produced amorphadiene (Fig. 5).

**EXAMPLE 2**

Isolation and characterization of the amorphadiene synthase encoding gene

A. Induction of transcription

As revealed in part III of figure 2, DHAA is photo-oxidatively converted into DHAA-OOH. In this reaction a reactive form of oxygen (singlet $O_2$) is added to DHAA. DHAA plays the role of an anti-oxidant, a scavenger of reactive oxygen species. Artemisinin is the stable end product of this reaction in which reactive oxygen is stored. Under stress conditions, (for example photo-stress, frost, drought or mechanical damage) reactive species of oxygen are formed in the plant. In response to this reactive oxygen generally plants are producing anti-oxidants. It is likely that A.annua will produce DHAA as anti-oxidant in response to this release of reactive oxygen. By exposing A.annua to stress conditions the transcription of the gene encoding amorphadiene synthase will be induced. To achieve this situation A.annua plants grown under climate room conditions (23°C, 90% moisture, 3000 lux) were exposed to stress conditions by putting them for one hour at approximately 30% moisture (drought stress) and 6000 lux (photo stress) at 30°C.

B. Isolation of total RNA

Total RNA of stress induced plants (according to example 2.A) was isolated from young leaves by the method of Verwoerd et al. (Nucleic Acids Research, **17**, 6, 2362 (1989)). DNase I (Deoxyribonuclease I, RNase free) was used to remove DNA from the RNA isolate. The DNase I was inactivated by exposure at 70°C during 15 minutes.

C. cDNA synthesis

The reverse transcription reaction was carried out in a 20 µl reaction containing 5 µg total RNA, 0.2 µg oligo $(dT)_{12}$, 0.5 mM each dATP, dTTP, dCTP and dGTP, 10 mM DTT, 2 U ribonuclease inhibitor (Gibco BRL), first strand synthesis buffer (Promega) and catalyzed with 200 U Moloney murine leukemia virus (M-MLV) reverse transcriptase RNase H minus (Promega). After 1 h incubation at 37°C the reaction was stopped by storing the reaction mixture at -20°C.

D. PCR-based probe generation

Based on comparison of sequences of terpenoid synthases, two degenerated primers were designed for two conserved regions. The sequence of the sense primer (primer A) was 5'-GA(C/T) GA(G/A) AA(C/T) GGI AA(G/A) TT(C/T) AA(G/A) GA-3' and the sequence of the anti sense primer (primer B) was 5'-CC (G/A)TA IGC (G/A)TC (G/A)AA IGT (G/A)TC (G/A)TC-3'. PCR was performed in a total volume of 100 µl containing 0.5 µM of each of these two primers, 0.2 mM each dNTP, 1 U Super Taq polymerase / 1x PCR buffer (HT Biotechnology LTD, Cambridge, England) and 2 µl cDNA. The reaction was incubated in a thermocycler (PTC 150, MJ-research) with 1 minute denaturation at 95°C, 1 minute annealing at 40°C and 1 minute and 15 seconds elongation at 72°C during 40 cycles. Agarose gel electrophoresis revealed a single specific PCR product of approximately 550 bp (538bp). Such a specific amplification product was only obtained when using cDNA made of RNA isolated from stress induced plants. The PCR

product was made blunt by using DNA polymerase I large fragment (Klenow), gel-purified and subcloned in Sma I digested pGEM 7Zf(+) (Stratagene) (Fig.6) and E.coli DH5α (Gibco BRL) was transformed with this construct. The inserts of 8 individual transformants were sequenced and they all had the same sequence as shown in figure 7.

E. cDNA Library construction

Synthesis of the second strand of the cDNA was done analogous to the RiboClone® cDNA synthesis System (Promega). After ligation with EcoR I (Not I) adapters (Gibco BRL) with sequence:

5'-pGTCGACGCGGCCGCG-3'
3'- CAGCTGCGCCGGCGCTTAA-OH-5'

the double stranded DNA was ligated into λ ExCell EcoR I/CIP (Pharmacia Biotech). For packaging and plating of the cDNA library, the Ready-To-Go® Lambda Packaging Kit (Pharmacia Biotech) was used. The titer of the unamplified library was $1.2 \times 10^6$ plaque forming units.

F. Library screening

For library screening 200 ng of the PCR amplified probe (Fig. 7) was gel purified, randomly labeled with [α-$^{32}$P]dCTP, according to the manufacturer's recommendation (Random Primed DNA Labeling Kit, Boehringer Mannheim Biochemica) and used to screen replica filters of $10^4$ plaques of the cDNA library plated on E.coli NM 522. The hybridization was performed for 16 h at 68°C in 1 M NaCl, 1% SDS and 10 % PEG (5000-7000). Filters were washed two times for 10 minutes at 50°C in 2 x SSC with 0.1 % SDS and exposed for 16 h to a Fuji X-ray film at -70°C. Clones yielding positive signals were isolated through a second and third round of hybridization. By transfecting E.coli NP66 (Pharmacia Biotech) with the positive clones, plasmid releases (Fig. 8) were obtained according to the manufacturer's

instructions (Pharmacia Biotech). Sequencing of these positive clones yielded a sequence as revealed in figure 8A.

**EXAMPLE 3**

Expression of the amorphadiene synthase encoding gene in E.coli BL21(DE3)

For functional expression the cDNA clone was subcloned in frame into the expression vector pET 11d (Stratagene). To introduce suitable restriction sites for subcloning, the gene was amplified by PCR using a sense primer (primer C) 5'-GT CGA CAA ACC **ATG** GCA CTT ACA GAA G-3' (introducing a NcoI site at the start codon **ATC**) and an anti-sense primer (primer D) 5'-GGATGGATCC **TCA** TAT ACT CAT AGG ATA AAC G-3' (introducing a BamHI site directly behind the stop codon **TGA**). The PCR reaction was performed under standard conditions. After digestion with BamHI and NcoI, the PCR product (Fig. 9[A]) and the expression vector pET 11d were gel purified and ligated together to yield a construct as revealed in figure 9.

To obtain expression; this gene construct (fig. 9), pET 11d without an insert as negative control, and pET 11d with the tobacco 5-epi-aristolochene synthase (TEAS) gene (Back et al., Archives of Biochemistry and Biophysics, **315**, 2, 527-532 (1994); Facchini & Chappell, Proc. Natl. Acad. Sci. USA, **89**, 11088-11092 (1992); Back & Chappell, The Journal of Biological Chemistry, 270, **13**, 7375-7381 (1995)) as positive control were transformed to E.coli BL21(DE3) (Stratagene), and grown overnight on LB agar plates supplemented with ampicillin at 37°C. Cultures of 50 ml LB medium supplemented with ampicillin (100 μg/ml) and 0.25 mM isopropyl-1-thio-ß-D-galactopyranoside (IPTG) were occulated with these over night cultures to $A_{600}$ = 0.5 and grown for 3 h at 27°C. The cells were harvested by centrifugation during 8 minutes at 2000 g and resuspended in 2 ml assay buffer. An aliquot of 1 ml resuspended cells was sonicated on ice four times for 5 seconds with 30 second intervals,

centrifuged for 5 minutes at 4°C in a microfuge (13.000 rpm) and the supernatant used for cyclase enzyme activity determinations and SDS-PAGE gel electroforese. Expression of the amorphadiene synthase gene-pET 11d construct (Fig. 9) in E.coli BL21(DE3) yielded a protein of approximately 50 to 60 kDa as shown in figure 10 lane 5 to 10. This agrees well to the size of amorphadiene synthase isolated from A.annua, which was determined to be 56 kDa (Fig. 13).

**EXAMPLE 4**

Conversion of FPP into amorphadiene by amorphadiene synthase expressed in E.coli.

Besides the supernatant of sonicated cells, also intact cells were used in the FPP assay. The FPP assay, GC-RAGA and GC-MS analyses were performed as described previously. Figure 11 and 11A are revealing the GC-RAGA chromatograms of the assays with intact transformed cells and with the supernatant of sonicated transformed cells, respectively. In both assays amorphadiene was produced. Identification of these assay products with the GC-MS gave a mass-spectrum identical to the mass-spectrum of the reference amorphadiene with a quality score of 99 % (maximum score), mass spectra were identical to the spectra as shown in Fig. 5. No amorphadiene was found in assays done with the positive and negative controls.

**EXAMPLE 5**

Expression of amorpha-4,11-diene synthase in transgenic tobacco

There are many ways to introduce DNA into a plant cell. Suitable methods by which DNA can be introduced into the plant cell include Agrobacterium infection or direct delivery of DNA such as, for example, by PEG-mediated transformation of protoplasts (Omirulleh et al., Plant Molecular Biology, **21**, 415-428 (1993)) or electroporation, by acceleration of DNA coated

microprojectiles (for example, microprojectile bombardment) microinjection, etc.

Because Agrobacterium tumefaciens-mediated transformation of Artemisia annua and Nicotiana tabacum with a sesquiterpene cyclase gene is known in literature (Vergauwe et al., Plant Cell Reports, **15**, 929-933 (1996); Hohn and Ohlrogge, Plant Physiol., **97**, 460-462 (1991)) delivery of expression units (cassettes), containing the amorphadiene synthase encoding gene, mediated by Agrobacterium seemed to be a rational approach.

There are several binary vector systems suitable to transfer the amorphadiene synthase encoding gene assembled in an expression cassette behind a suitable promoter (for example, the cauliflower mosaic virus 35S promoter) and upstream of a suitable terminator (for example, the nopaline synthase transcription terminator) to tobacco and/or A.annua. To mobilize the recombinant binary vector to the target Agrobacterium strain, a triparental mating procedure was carried out by using E.coli carrying the recombinant binary vector and a helper E.coli strain carrying a plasmid (pRK2013 for example) which was able to mobilize the recombinant binary vector to the target Agrobacterium strain. This transformed Agrobacterium strain was used for transformation of explants from the target plant species. Only the transformed tissue carrying a resistance marker (for example kanamycin-resistance, present between the binary plasmid T-DNA borders) regenerated on a selectable (for example kanamycin containing) regeneration medium. (According to Rogers SG, Horsch RB, Fraley RT Methods Enzymol (1986)118: 627-640).

**EXAMPLE 6**

Conversion of amorphadiene into artemisinin (DHAA) by A.annua

This assay was carried out in a way analogous to the method as described by Koepp et al. (The Journal of Biological Chemistry, 270, **15**, 8686-8690 (1995)).

Radioactive ($^3$H-labeled) amorphadiene was fed to leaf discs of A.annua. For the infiltration of amorphadiene into the leaf discs of A.annua the radioactive amorphadiene can be made water soluble by complexation with cyclodextrins, for example. Radioactive amorphadiene is obtained by using the FPP-assay with the transformed E.coli BL21(DE3) cells (carrying the cloned amorphadiene synthetase gene of A.annua). Identification of the product(s) made in this assay was done by radio-GC analysis in combination with MS and or GC-MS analysis. The expected intermediates arteannuic acid (AA), dihydroarteannuic acid (DHAA) and the end product artemisinin are all used as references.

**EXAMPLE 7**

Expression of amorpha-4,11-diene synthase in transgenic A.annua and the production of artemisinin

Transformed A.annua plants were prepared as described in example 5.

For the regeneration of A.annua the medium for callus, shoot and root induction consisted of Murashige and Skoog micro and macro elements including modified vitamins (Duchefa Biochemie, Haarlem, The Netherlands), 4 % (w/v) sucrose, 0,1 mg/L Indole-3-acetic acid (IAA), 0,1 mg/L 6-benzylaminopurine (BAP) and 0,8 % (w/v) agar. The pH was adjusted to 5.7 with NaOH prior to the addition of agar. The medium was autoclaved at 1 bar for 20 min. Transformed explants were regenerated on this medium to fully regenerated plants.

## CLAIMS

1. Isolated DNA sequence encoding a polypeptide having the biological activity of amorpha-4,11-diene synthase.

2. DNA sequence as claimed in claim 1 which exhibits at least 70% homology to the sequence as shown in Figure 9A or the complementary strand thereof and which codes for a polypeptide having the biological activity of the enzyme amorphadiene synthase.

3. DNA sequence as claimed in claim 2, which is at least 80%, preferably at least 90%, more preferably at least 95% homologous to the sequence in Figure 9A.

4. DNA sequence as claimed in claims 1-3, which has the sequence as shown in Figure 9A.

5. DNA sequence as claimed in claims 1-4, characterized in that it has been isolated from plants producing amorpha-4,11-diene, for example A.annua and V.oblongifolia

6. Use of a DNA sequence as claimed in claims 1-5 for transforming or transfecting a host cell.

7. DNA construct comprising the DNA sequence as claimed in claims 1-5 operably linked to suitable transcription initiation and termination sequences.

8. Host cell comprising a DNA sequence as claimed in claims 1 to 5 or a DNA construct as claimed in claim 7.

9. Host cell as claimed in claim 8, wherein the cell is a bacterial cell, in particular an E.coli cell.

10. Host cell as claimed in claim 8, wherein the cell is a plant cell.

11. Host cell as claimed in claim 10, wherein the cell is derived from a plant itself producing sesquiterpenes.

12. Host cell as claimed in claim 11, wherein the cell is a A.annua cell or a V.oblongifolia cell.

13. Host cell as claimed in claim 11, wherein the cell is derived from a plant selected from the group

consisting of the genera Carum, Cichorium, Daucus, Juniperus, Chamomilla, Lactuca, Pogostemon and Vetiveria.

14. Host cell as claimed in claim 10, wherein the cell is derived from a plant in which the biosynthesis of sesquiterpenoids can be induced by elicitation.

15. Host cell as claimed in claim 14, wherein the cell is derived from a plant selected from the group consisting of the genera Capsicum, Gossypium, Lycopersicon, Nicotiana, Phleum, Solanum and Ulmus.

16. Host cell as claimed in claim 10, wherein the cell is derived from a plant selected from the group of soybean, sunflower and rapeseed.

17. Host cell as claimed in claims 8 and 10-16, which cell is part of a tissue or organism.

18. Transgenic tissue, consisting at least part of host cells as claimed in claims 8 and 10-16.

19. Transgenic organism, consisting at least part of host cells as claimed in claims 8 and 10-16.

20. Polypeptide having the biological activity of the enzyme amorphadiene synthase in isolated form obtainable by isolating the polypeptide from A.annua or V.oblongifolia by a process as described in Example 1.

21. Recombinant polypeptide having the biological activity of the enzyme amorphadiene synthase obtainable by expressing a DNA sequence as claimed in claims 1-5 in a suitable host cell as claimed in claims 8-17.

22. Method of preparing amorphadiene, comprising:

a) incubating a polypeptide as claimed in claim 20 and 21 in the presence of farnesyl pyrophosphate (FPP) in an incubation medium at a suitable temperature and during a suitable period of time; and

b) optionally isolating the amorphadiene thus formed.

23. Method of preparing amorphadiene, comprising the steps of:

a) transfecting or transforming a suitable host cell with a DNA sequence as claimed in claims 1-5 or a construct according to claim 7 to obtain transgenic host cells;

b) expressing the said DNA sequence in the presence of farnesyl pyrophosphate (FPP) to form amorphadiene; and

c) optionally isolating the amorphadiene thus formed,

wherein the expression level of the amorphadiene synthase is higher in transgenic host cells, tissues or organisms harboring an endogenous version of the DNA sequence than in non-transgenic host cells, tissues or organisms.

24. Method of preparing artemisinin, comprising:

a) incubation of a polypeptide as claimed in claim 20 and 21 in the presence of farnesyl pyrophosphate (FPP) and the enzymes that further convert amorpha-4,11-diene to artemisinin in an incubation medium at a suitable temperature and during a suitable period of time; and

b) optional isolation of the artemisinin thus formed.

25. Method of preparing artemisinin, comprising:

a) transfecting or transforming a suitable host cell, tissue or organism with a DNA sequence as claimed in claims 1-5 or a construct according to claim 7 to obtain transgenic host cells, tissues or organisms;

b) expressing the said DNA sequence in the presence of farnesyl pyrophosphate (FPP); and

c) optionally isolating the amorpha-4,11-diene thus formed,

wherein the transgenic host cells, tissues or organisms harbor the genetic information coding for the enzymes that further convert amorpha-4,11-diene to artemisinin and wherein the expression level of the amorpha-4,11-diene synthase is higher in transgenic host cells,

tissues or organisms harboring an endogenous version of the DNA sequence than in non-transgenic host cells, tissues or organisms.

26. Source of artemisinin, comprising host cells, tissues or organisms harboring a DNA sequence as claimed in claims 1-5 and the genetic information coding for the enzymes that further convert amorpha-4,11-diene to artemisinin, which host cells, tissues or organisms have expressed the said DNA sequence.

27. Source as claimed in claim 26, wherein the cells are bacterial cells or plant cells.

28. Source as claimed in claim 26, wherein the cells are disrupted.

29. Transgenic cell, tissue or organism harboring in its genome more copies of a DNA sequence as claimed in claims 1-5 than are present in a corresponding non-transgenic cell, tissue or organism.

30. Transgenic cell as claimed in claim 29, which cell is an E.coli cell.

31. Transgenic organism as claimed in claim 29, which organism is a plant itself producing sesquiterpenes.

32. Transgenic organism as claimed in claim 31, which organism is A.annua or V.oblongifolia.

33. Transgenic organism as claimed in claim 31, which organism is a plant selected from the group consisting of the genera Carum, Cichorium, Daucus, Juniperus, Chamomilla, Lactuca, Pogostemon and Vetiveria.

34. Transgenic organism as claimed in claim 29, which organism is a plant in which the biosynthesis of sesquiterpenoids can be induced by elicitation.

35. Transgenic organism as claimed in claim 34, which organism is a plant selected from the group consisting of the genera Capsicum, Gossypium, Lycopersicon, Nicotiana, Phleum, Solanum and Ulmus.

36. Transgenic organism as claimed in claim 29, which organism is a plant selected from the group consisting of soybean, sunflower and rapeseed.

EPO - DG 1

27. 08. 1998

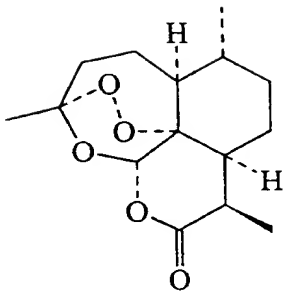



Fig. 1. Artemisinin

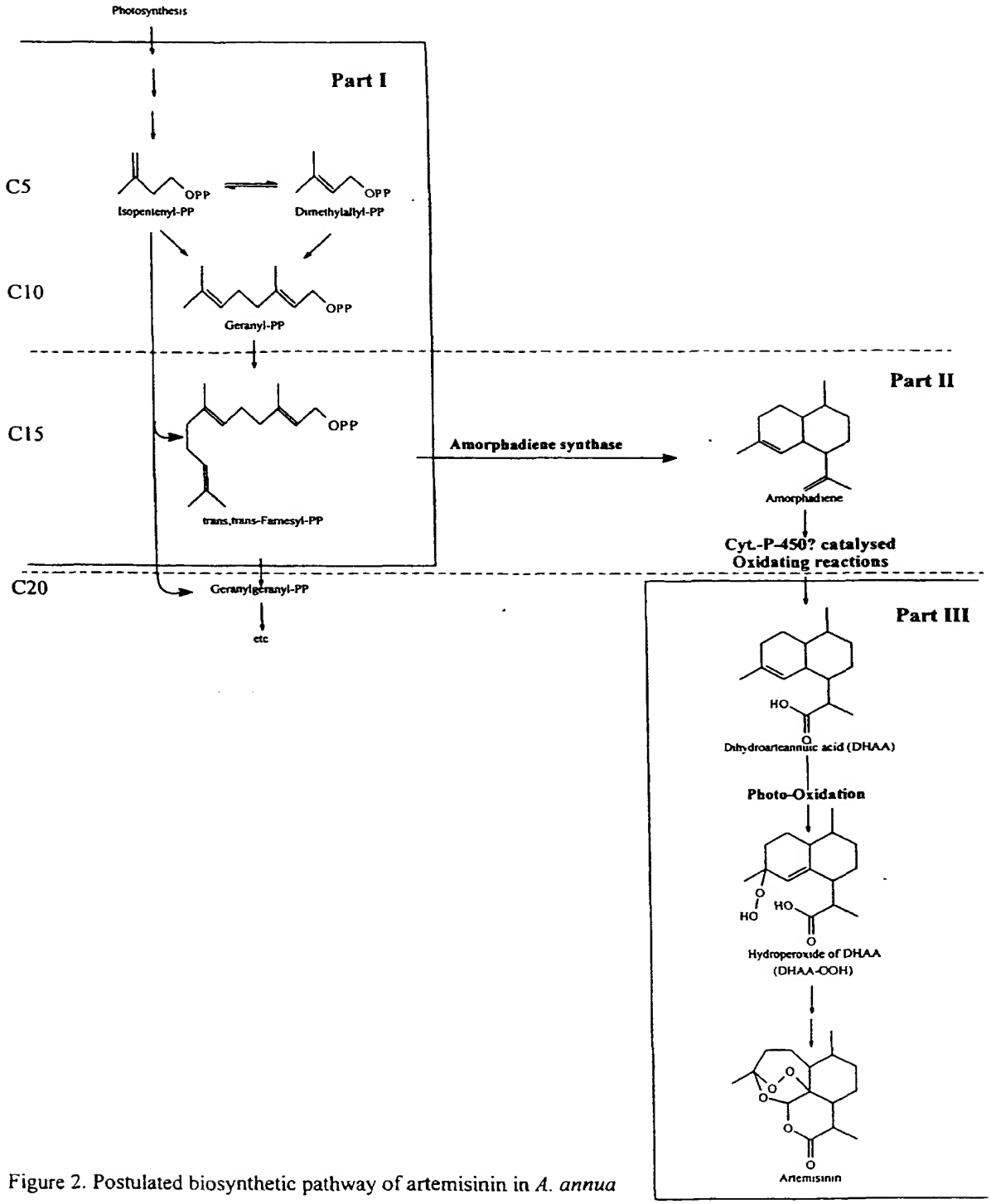



Figure 2. Postulated biosynthetic pathway of artemisinin in *A. annua*

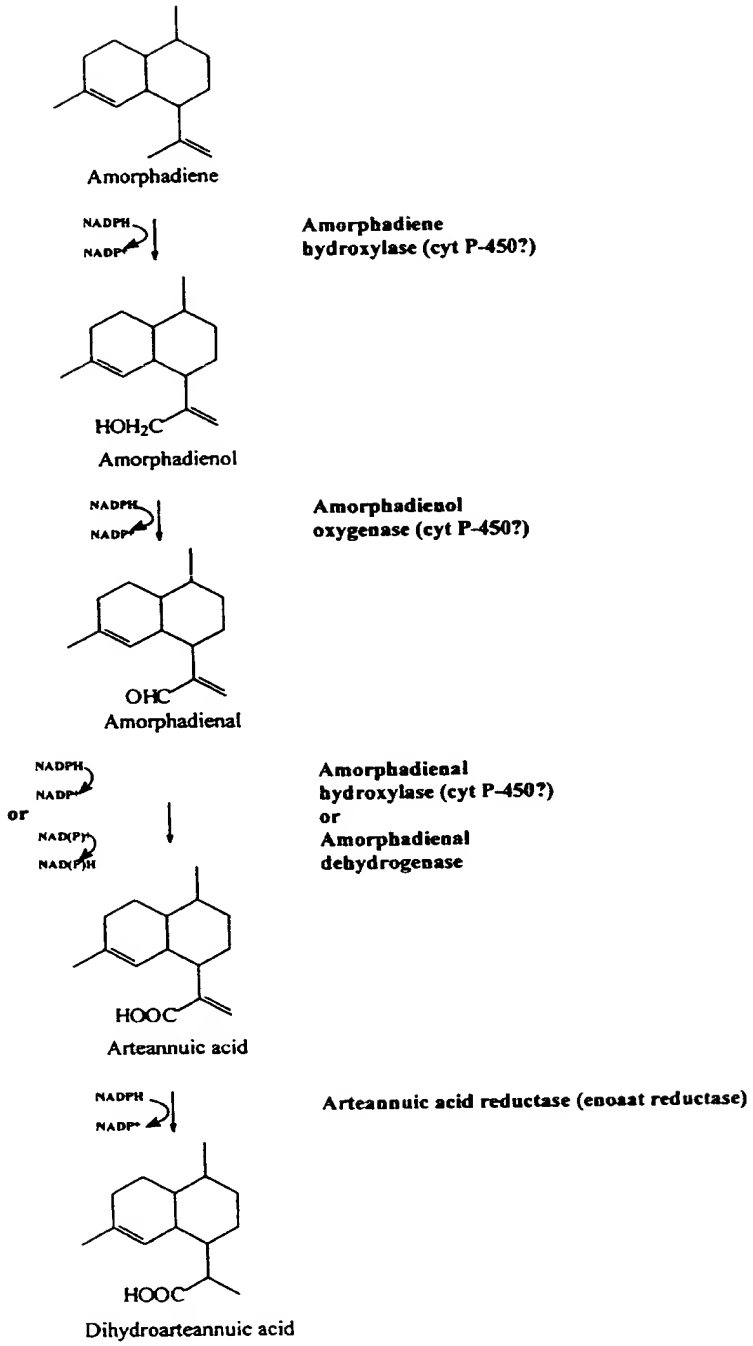



Figure 2A. Transition between part II and part III. Hypothetical conversion of amorphadiene into dihydroarteannuic acid in *A. annua*

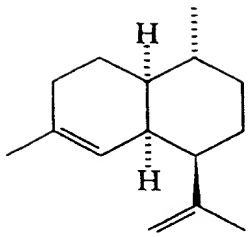



Fig. 3 Amorpha-4,11-diene

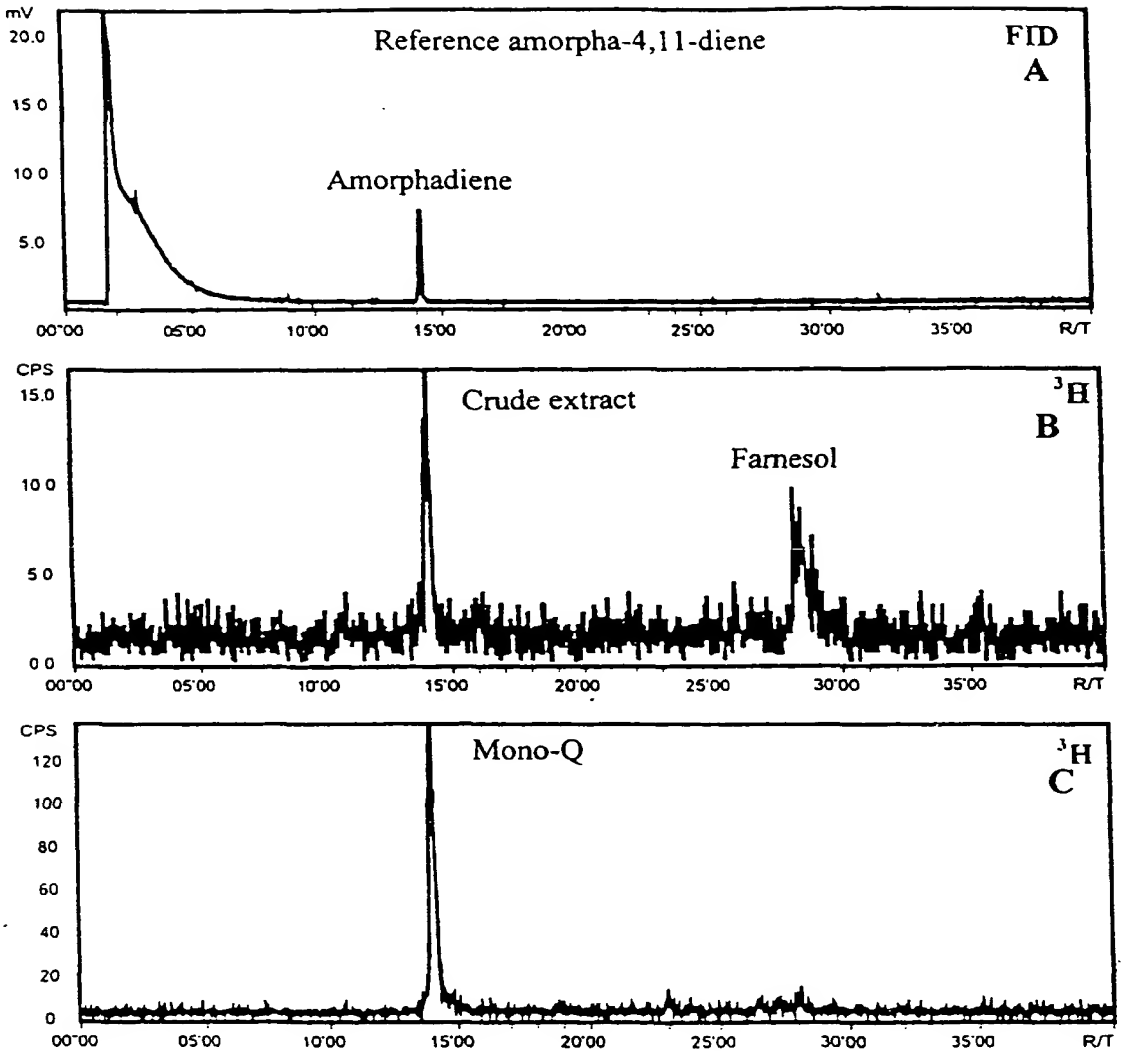



Fig. 4. Radio-GC chromatograms of the [$^3$H]-FPP-assays. **A**: Flame Ionization Detector (FID) signal of amorphadiene (reference). **B**: Radio signals of the $^3$H labeled assay products amorphadiene (ret. time 14 min.) and farnesol (as a product of aspecific phosphohydrolase activity, ret. time 28 min.) obtained with crude enzyme extract. **C**: Radio signal of the $^3$H labeled assay product amorphadiene obtained with Mono-Q purified enzyme extract.

Library Searched : C:\DATABASE\WITLOF.L
Quality : 99
ID : amorpha-4,11-dieen

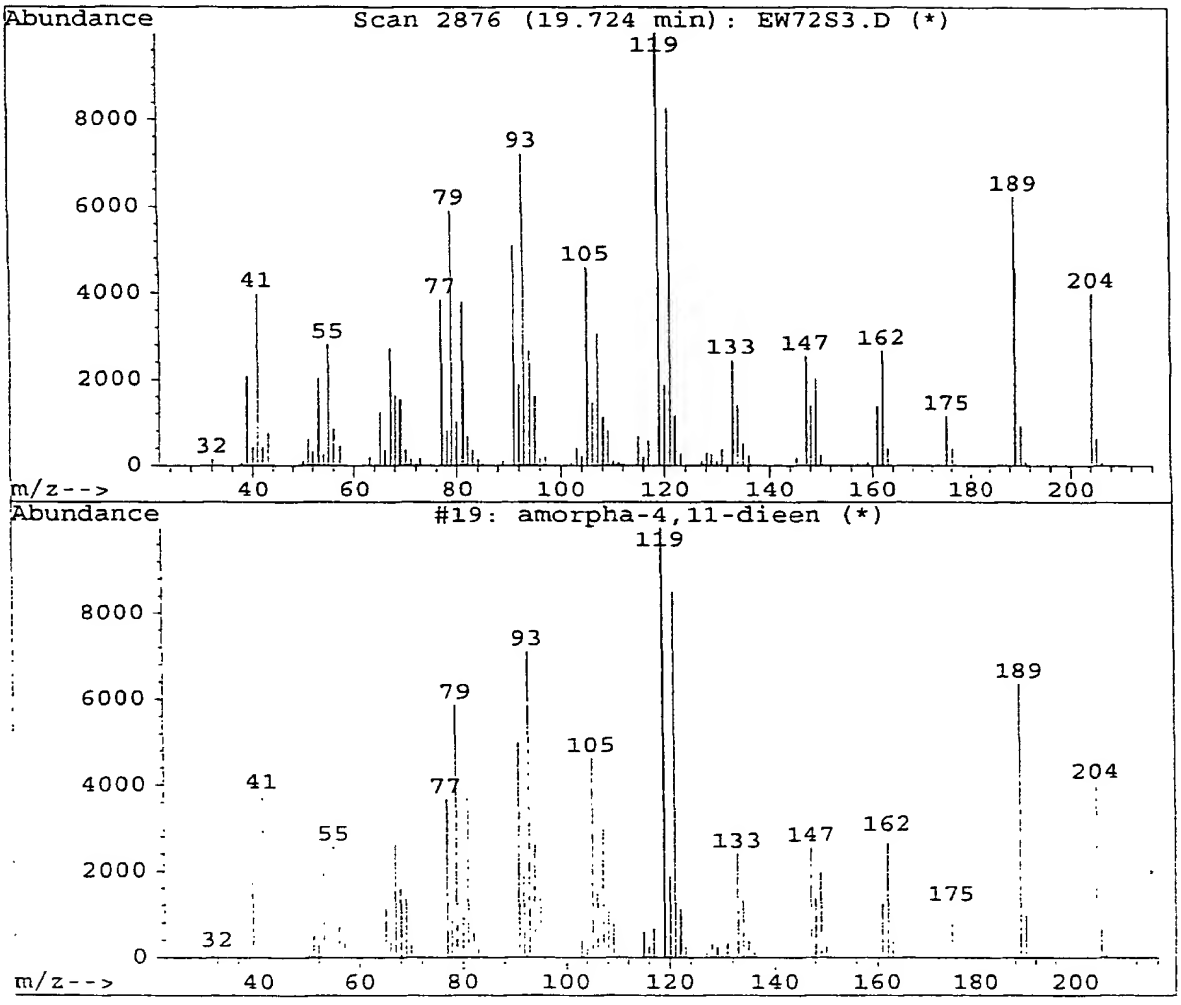



Fig. 5. Mass spectrum of reference amorphadiene compared with the mass spectrum of the FPP assay with terpene cyclases (synthases) purified from *A. annua*. This comparison yielded a quality score of 99 %, corresponding with a maximum score of identicalness.

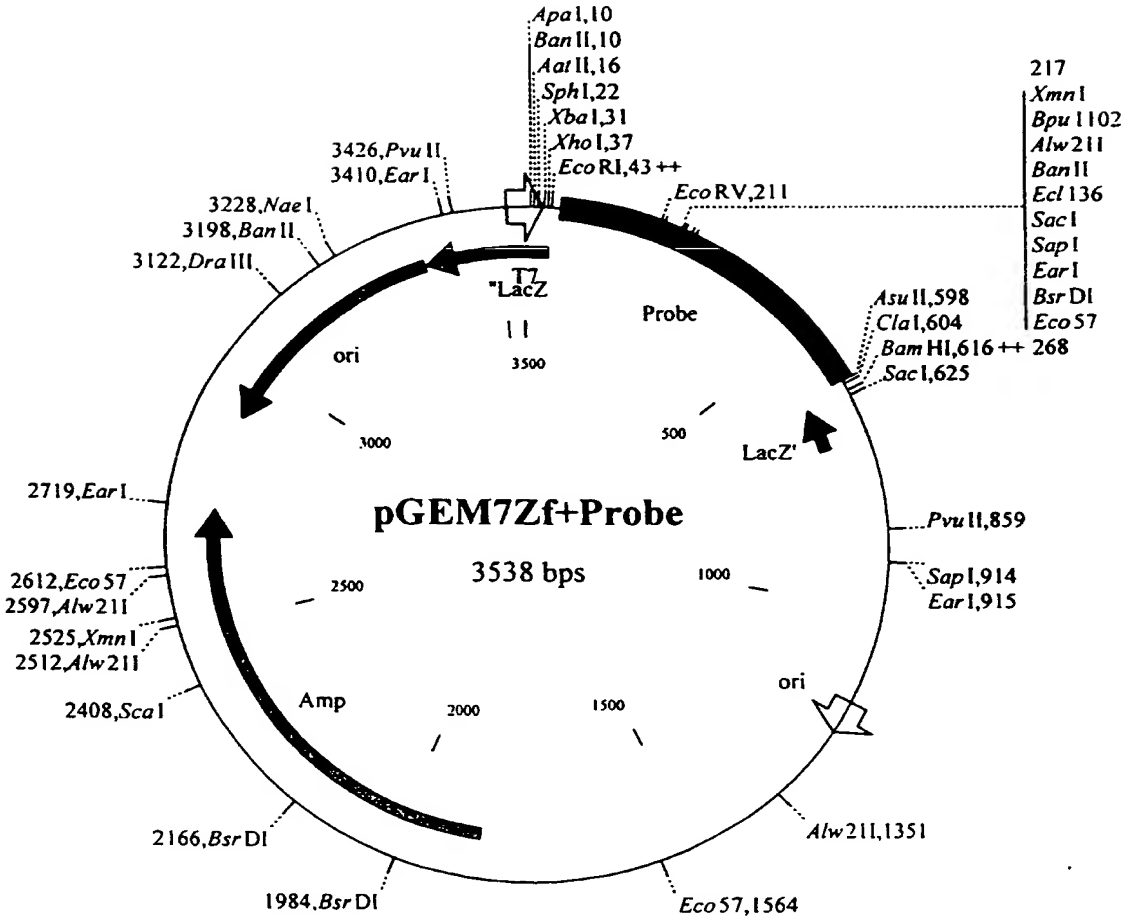



Fig. 6. Probe generated by PCR and cloned into pGEM 7Zf*.

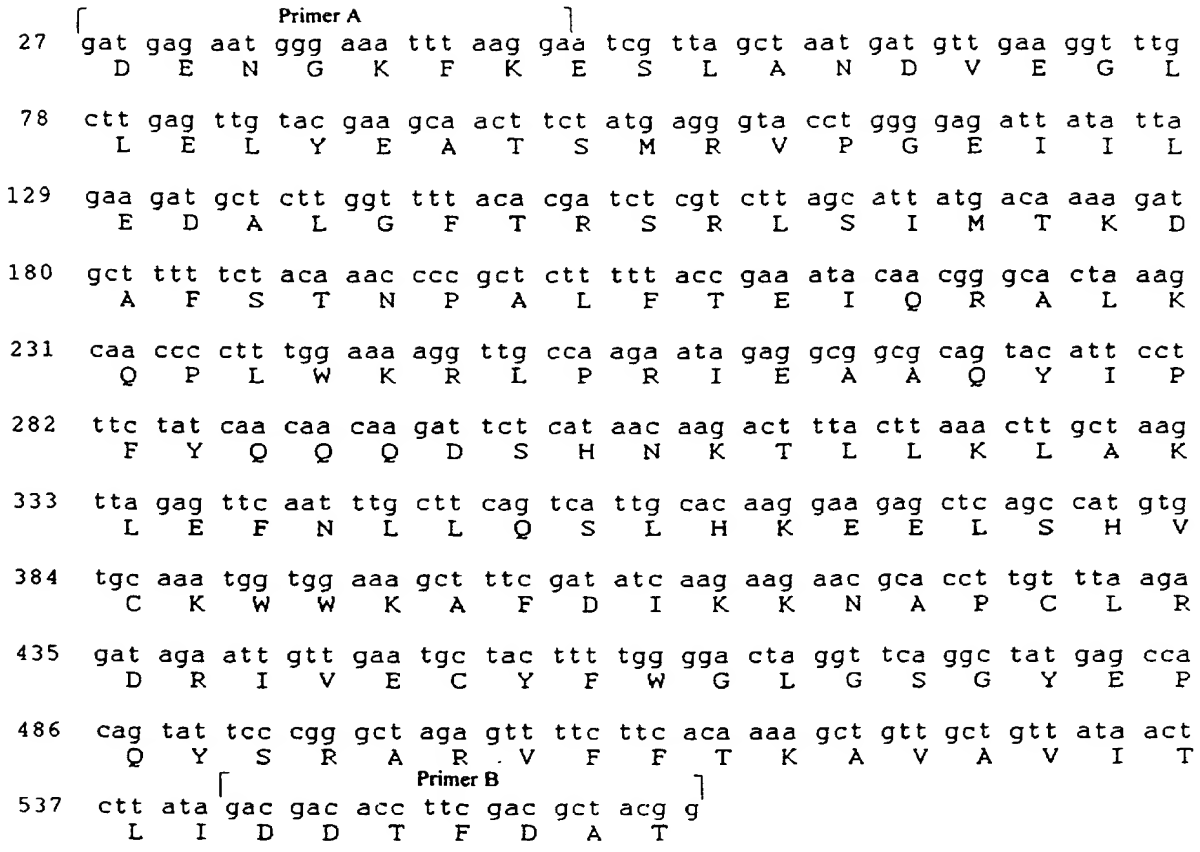



Fig. 7. Nucleotide acid sequence and deduced amino acid sequence of the probe (538 bp) generated by PCR with the primers A and B.

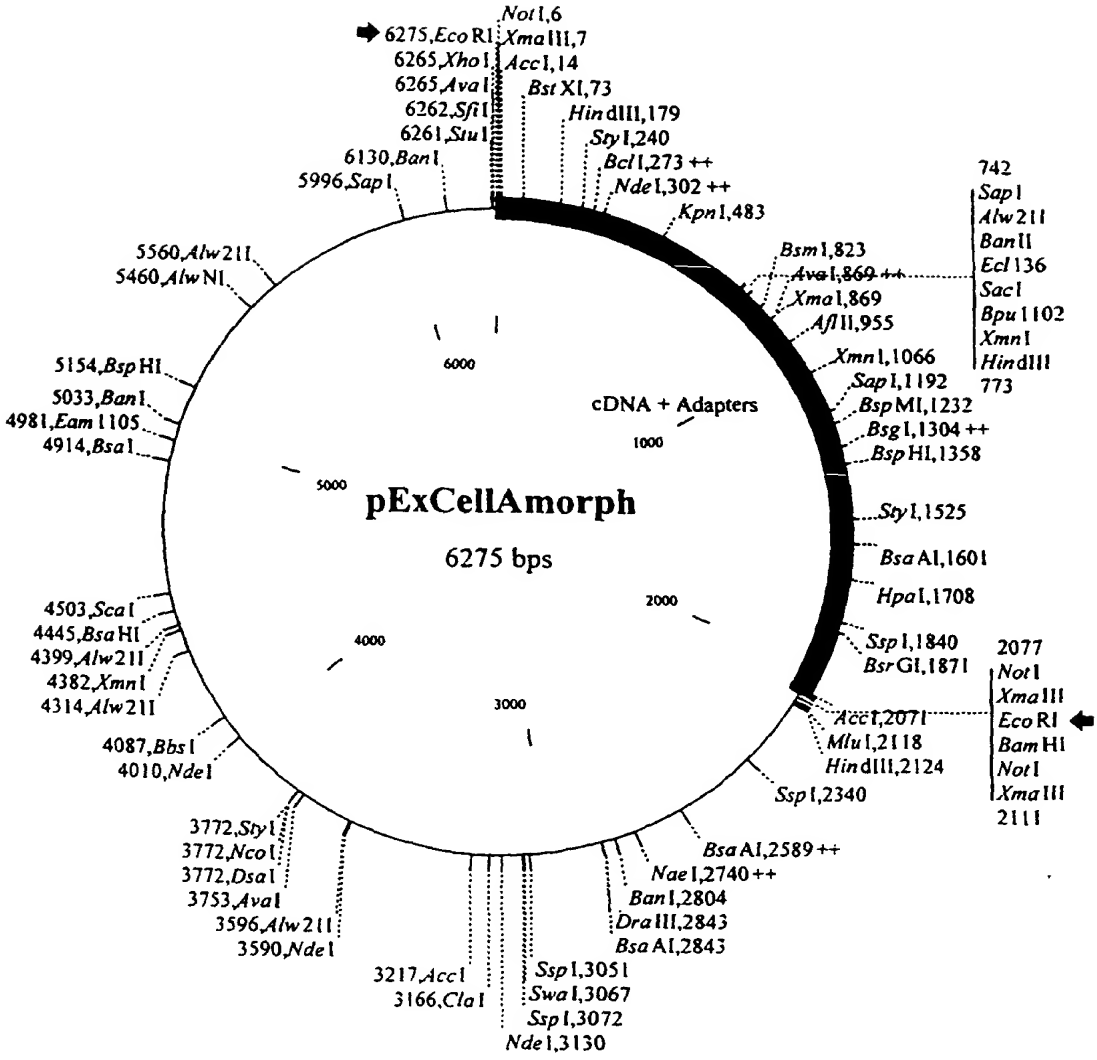
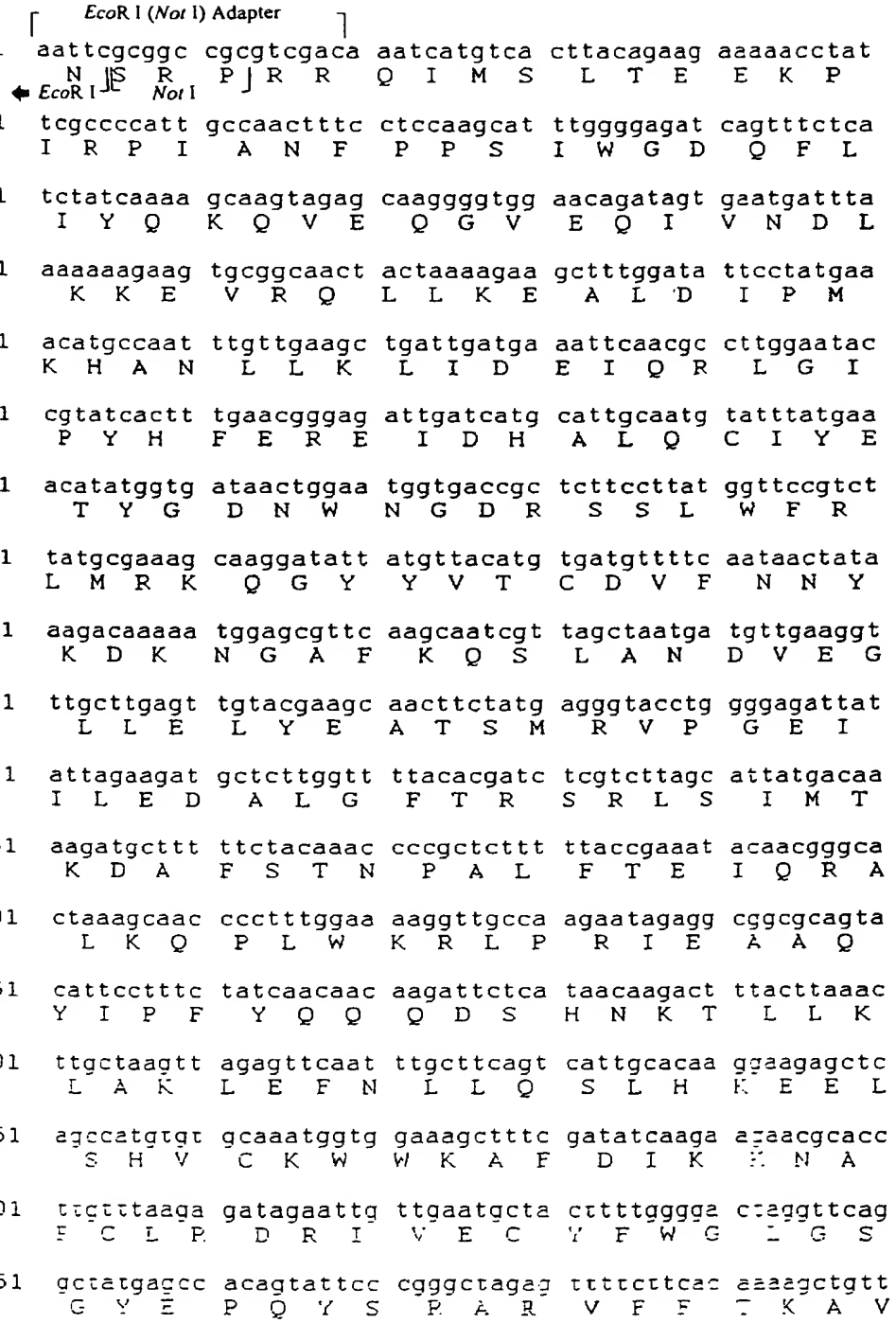



Fig. 8. Released plasmid of a positive clone isolated from the cDNA library of induced *A. annua*.




```
    aattcgcggc cgcgtcgaca aatcatgtca cttacagaag aaaaacctat
      N   S   R   P   R   R   Q   I   M   S   L   T   E   E   K   P
    tcgccccatt gccaactttc ctccaagcat ttggggagat cagtttctca
      I   R   P   I   A   N   F   P   P   S   I   W   G   D   Q   F   L
    tctatcaaaa gcaagtagag caaggggtgg aacagatagt gaatgattta
      I   Y   Q   K   Q   V   E   Q   G   V   E   Q   I   V   N   D   L
    aaaaaagaag tgcggcaact actaaaagaa gctttggata ttcctatgaa
      K   K   E   V   R   Q   L   L   K   E   A   L   D   I   P   M
    acatgccaat ttgttgaagc tgattgatga aattcaacgc cttggaatac
      K   H   A   N   L   L   K   L   I   D   E   I   Q   R   L   G   I
    cgtatcactt tgaacgggag attgatcatg cattgcaatg tatttatgaa
      P   Y   H   F   E   R   E   I   D   H   A   L   Q   C   I   Y   E
    acatatggtg ataactggaa tggtgaccgc tcttccttat ggttccgtct
      T   Y   G   D   N   W   N   G   D   R   S   S   L   W   F   R
    tatgcgaaag caaggatatt atgttacatg tgatgttttc aataactata
      L   M   R   K   Q   G   Y   Y   V   T   C   D   V   F   N   N   Y
    aagacaaaaa tggagcgttc aagcaatcgt tagctaatga tgttgaaggt
      K   D   K   N   G   A   F   K   Q   S   L   A   N   D   V   E   G
    ttgcttgagt tgtacgaagc aacttctatg agggtacctg gggagattat
      L   L   E   L   Y   E   A   T   S   M   R   V   P   G   E   I
    attagaagat gctcttggtt ttacacgatc tcgtcttagc attatgacaa
      I   L   E   D   A   L   G   F   T   R   S   R   L   S   I   M   T
    aagatgcttt ttctacaaac cccgctcttt ttaccgaaat acaacgggca
      K   D   A   F   S   T   N   P   A   L   F   T   E   I   Q   R   A
    ctaaagcaac ccctttggaa aaggttgcca agaatagagg cggcgcagta
      L   K   Q   P   L   W   K   R   L   P   R   I   E   A   A   Q
    cattcctttc tatcaacaac aagattctca taacaagact ttacttaaac
      Y   I   P   F   Y   Q   Q   Q   D   S   H   N   K   T   L   L   K
    ttgctaagtt agagttcaat ttgcttcagt cattgcacaa ggaagagctc
      L   A   K   L   E   F   N   L   L   Q   S   L   H   K   E   E   L
    agccatgtgt gcaaatggtg gaaagctttc gatatcaaga agaacgcacc
      S   H   V   C   K   W   W   K   A   F   D   I   K   K   N   A
    ttgtttaaga gatagaattg ttgaatgcta cttttgggga ctaggttcag
      F   C   L   R   D   R   I   V   E   C   Y   F   W   G   L   G   S
    gctatgagcc acagtattcc cgggctagag ttttcttcac aaaagctgtt
      G   Y   E   P   Q   Y   S   R   A   R   V   F   F   T   K   A   V
```

```
901  gctgttataa ctcttataga tgacacttat gatgcgtatg gtacttatga
       A  V  I     T  L  I     D  D  T  Y     D  A  Y     G  T  Y

951  agaacttaag atctttactg aagctgttga aaggtggtca attacatgct
     E  E  L  K     I  F  T     E  A  V     E  R  W  S     I  T  C

1001 tagacacact tccagaatac atgaaaccga tatacaaatt attcatggat
       L  D  T     L  P  E  Y     M  K  P     I  Y  K     L  F  M  D

1051 acatacacag aaatggaaga atttcttgca aaggagggaa gaacagatct
       T  Y  T     E  M  E     E  F  L  A     K  E  G     R  T  D

1101 atttaactgc ggcaaagaat ttgtgaaaga gtttgttaga aacctgatgg
     L  F  N  C     G  K  E     F  V  K     E  F  V  R     N  L  M

1151 ttgaagcaaa atgggcaaat gagggacaca taccaaccac tgaagagcat
       V  E  A     K  W  A  N     E  G  H     I  P  T     T  E  E  H

1201 gatccagttg taatcattac tggcggtgct aacctgctta caacaacttg
       D  P  V     V  I  I     T  G  G  A     N  L  L     T  T  T

1251 ttatcttggc atgagtgata tattcacaaa agagtctgtc gaatgggctg
     C  Y  L  G     M  S  D     I  F  T     K  E  S  V     E  W  A

1301 tctctgcacc tcctcttttt agatactcag gtatacttgg tcgacgccta
       V  S  A     P  P  L  F     R  Y  S     G  I  L     G  R  R  L

1351 aatgatctca tgacccacaa ggccgagcaa gaaagaaaac atagttcatc
       N  D  L     M  T  H     K  A  E  Q     E  R  K     H  S  S

1401 gagccttgaa agttatatga aggaatataa tgtcaatgag gagtatgccc
     S  S  L  E     S  Y  M     K  E  Y     N  V  N  E     E  Y  A

1451 aaaccttgat ttacaaggaa gtagaagatg tgtggaaaga tataaaccga
       Q  T  L     I  Y  K  E     V  E  D     V  W  K     D  I  N  R

1501 gagtacctca caactaaaaa cattccaagg ccgttattga tggctgtgat
       E  Y  L     T  T  K     N  I  P  R     P  L  L     M  A  V

1551 ctatttgtgc cagtttcttg aagttcaata tgcaggaaag gataacttca
     I  Y  L  C     Q  F  L     E  V  Q     Y  A  G  K     D  N  F

1601 cacgtatggg agacgaatac aaacatctca taaagtctct actcgtttat
       T  R  M     G  D  E  Y     K  H  L     I  K  S     L  L  V  Y

1651 cctatgagta tatgactacc aatccttcgt gcatagccta tcaattatat
       P  M  S     I  -  L     P  I  L  R     A  -  P     I  N  Y

1701 tgaaagggtt aactatgcac gtctctatgg agagaatttc tcaagctatt
     I  E  R  V     N  Y  A     R  L  Y     G  E  N  F     S  S  Y

1751 tggtgtttct tgctggcaat aataaatcag acgcataaaa ttgtattgaa
     L  V  F     L  A  G  N     N  K  S     D  A  -     N  C  I  E

1801 ctatatgccg atagctattt aaagttatta tacaactaaa atattcaaca
       L  Y  A     D  S  Y     L  K  L  L     Y  N  -     N  I  Q

1851 atggtattat acttttactt tgtacaaaag caaaagtaca ctactgttat
     Q  W  Y  Y     T  F  T     L  Y  K     S  K  S  T     L  L  L
```

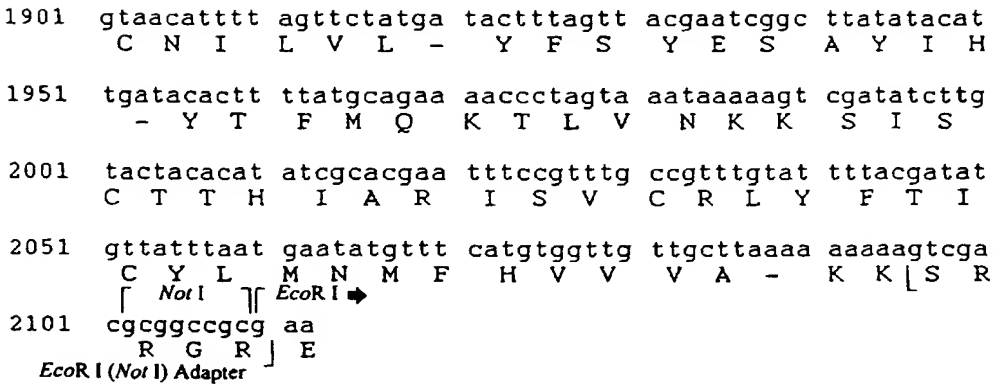



Fig. 8[A]. Nucleotide acid sequence and deduced amino acid sequence of a positive clone (amorphadiene synthase encoding gene) isolated from the cDNA library of induced *A. annua*. The sequence is flanked with *Eco*R I (*Not* I) adapters (Gibco BRL).

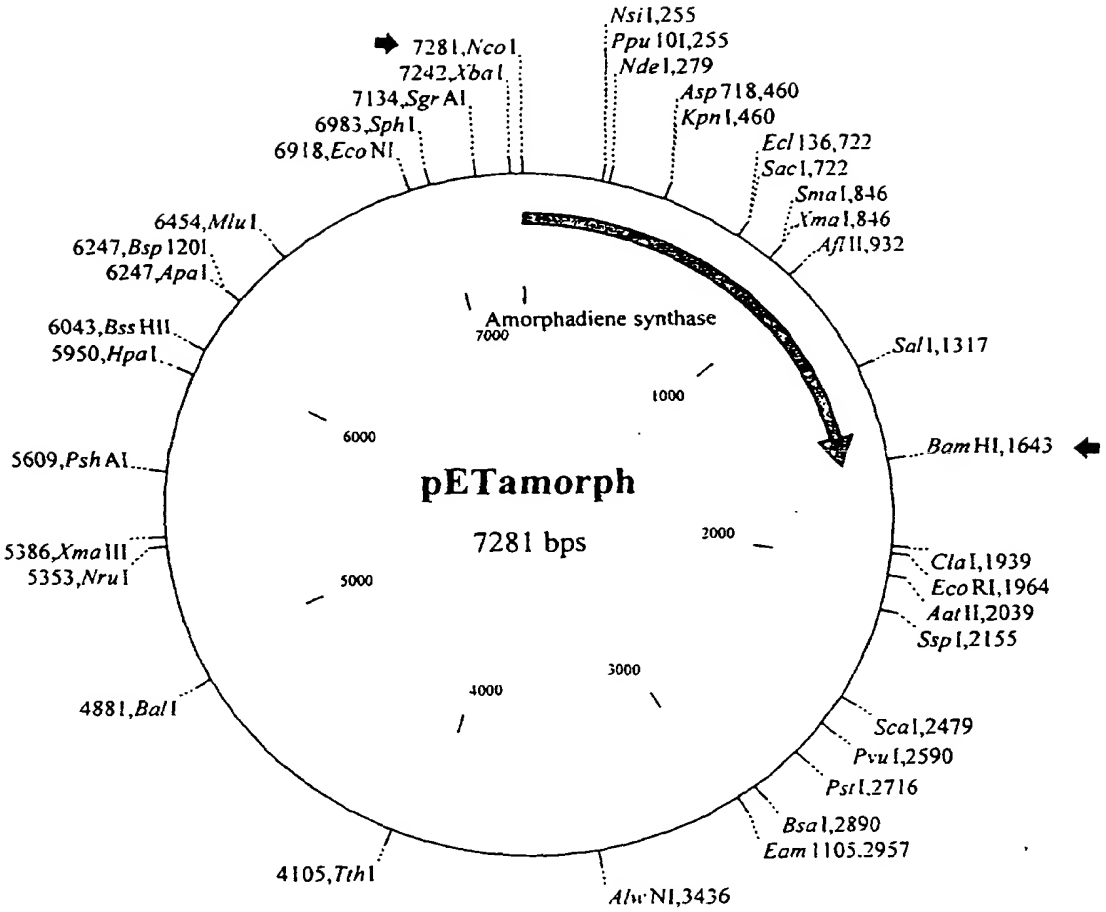
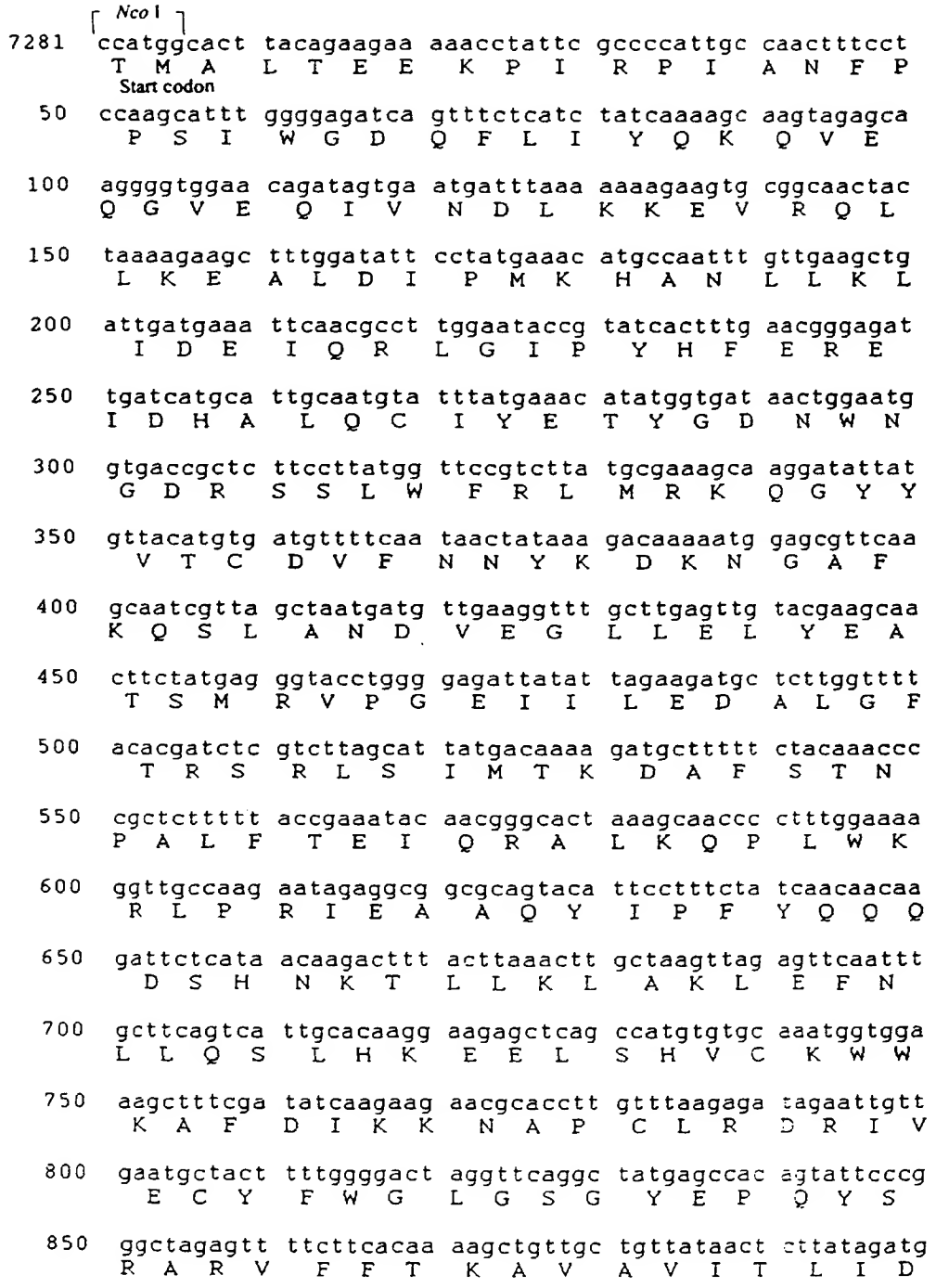



Fig. 9. Part, between start and stop codon(flanked resp. with a *Nco* I and *Bam*H I site), of the amorphadiene synthase encoding gene cloned (*Nco* I, *Bam*H I) in the expression vector pET 11d.


Nco I
7281 ccatggcact tacagaagaa aaacctattc gccccattgc caactttcct
T M A L T E E K P I R P I A N F P
Start codon
50 ccaagcattt ggggagatca gtttctcatc tatcaaaagc aagtagagca
P S I W G D Q F L I Y Q K Q V E
100 aggggtggaa cagatagtga atgatttaaa aaaagaagtg cggcaactac
Q G V E Q I V N D L K K E V R Q L
150 taaaagaagc tttggatatt cctatgaaac atgccaattt gttgaagctg
L K E A L D I P M K H A N L L K L
200 attgatgaaa ttcaacgcct tggaataccg tatcactttg aacgggagat
I D E I Q R L G I P Y H F E R E
250 tgatcatgca ttgcaatgta tttatgaaac atatggtgat aactggaatg
I D H A L Q C I Y E T Y G D N W N
300 gtgaccgctc ttccttatgg ttccgtctta tgcgaaagca aggatattat
G D R S S L W F R L M R K Q G Y Y
350 gttacatgtg atgttttcaa taactataaa gacaaaaatg gagcgttcaa
V T C D V F N N Y K D K N G A F
400 gcaatcgtta gctaatgatg ttgaaggttt gcttgagttg tacgaagcaa
K Q S L A N D V E G L L E L Y E A
450 cttctatgag ggtacctggg gagattatat tagaagatgc tcttggtttt
T S M R V P G E I I L E D A L G F
500 acacgatctc gtcttagcat tatgacaaaa gatgcttttt ctacaaaccc
T R S R L S I M T K D A F S T N
550 cgctcttttt accgaaatac aacgggcact aaagcaaccc ctttggaaaa
P A L F T E I Q R A L K Q P L W K
600 ggttgccaag aatagaggcg gcgcagtaca ttcctttcta tcaacaacaa
R L P R I E A A Q Y I P F Y Q Q Q
650 gattctcata acaagacttt acttaaactt gctaagttag agttcaattt
D S H N K T L L K L A K L E F N
700 gcttcagtca ttgcacaagg aagagctcag ccatgtgtgc aaatggtgga
L L Q S L H K E E L S H V C K W W
750 aagctttcga tatcaagaag aacgcacctt gtttaagaga tagaattgtt
K A F D I K K N A P C L R D R I V
800 gaatgctact tttggggact aggttcaggc tatgagccac agtattcccg
E C Y F W G L G S G Y E P Q Y S
850 ggctagagtt ttcttcacaa aagctgttgc tgttataact cttatagatg
R A R V F F T K A V A V I T L I D

```
900  acacttatga tgcgtatggt acttatgaag aacttaagat ctttactgaa
       D  T  Y   D  A  Y  G   T  Y  E   E  L  K   I  F  T  E

950  gctgttgaaa ggtggtcaat tacatgctta gacacacttc cagaatacat
       A  V  E   R  W  S   I  T  C  L   D  T  L   P  E  Y

1000 gaaaccgata tacaaattat tcatggatac atacacagaa atggaagaat
     M  K  P  I   Y  K  L   F  M  D   T  Y  T  E   M  E  E

1050 ttcttgcaaa ggagggaaga acagatctat ttaactgcgg caaagaattt
       F  L  A   K  E  G  R   T  D  L   F  N  C   G  K  E  F

1100 gtgaaagagt ttgttagaaa cctgatggtt gaagcaaaat gggcaaatga
       V  K  E   F  V  R   N  L  M  V   E  A  K   W  A  N

1150 gggacacata ccaaccactg aagagcatga tccagttgta atcattactg
     E  G  H  I   P  T  T   E  E  H   D  P  V  V   I  I  T

1200 gcggtgctaa cctgcttaca acaacttgtt atcttggcat gagtgatata
       G  G  A   N  L  L  T   T  T  C   Y  L  G   M  S  D  I

1250 ttcacaaaag agtctgtcga atgggctgtc tctgcacctc ctcttttag
       F  T  K   E  S  V   E  W  A  V   S  A  P   P  L  F

1300 atactcaggt atacttggtc gacgcctaaa tgatctcatg acccacaagg
     R  Y  S  G   I  L  G   R  R  L   N  D  L  M   T  H  K

1350 ccgagcaaga aagaaaacat agttcatcga gccttgaaag ttatatgaag
       A  E  Q   E  R  K  H   S  S  S   S  L  E   S  Y  M  K

1400 gaatataatg tcaatgagga gtatgcccaa accttgattt acaaggaagt
       E  Y  N   V  N  E   E  Y  A  Q   T  L  I   Y  K  E

1450 agaagatgtg tggaaagata taaaccgaga gtacctcaca actaaaaaca
     V  E  D  V   W  K  D   I  N  R   E  Y  L  T   T  K  N

1500 ttccaaggcc gttattgatg gctgtgatct atttgtgcca gtttcttgaa
       I  P  R   P  L  L  M   A  V  I   Y  L  C   Q  F  L  E

1550 gttcaatatg caggaaagga taacttcaca cgtatgggag acgaatacaa
       V  Q  Y   A  G  K   D  N  F  T   R  M  G   D  E  Y
                                                  ┌ BamH I ┐
1600 acatctcata aagtctctac tcgtttatcc tatgagtata tgaggatcc
     K  H  L  I   K  S  L   L  V  Y   P  M  S  I   -  G  S
                                          Stop codon
```

Fig. 9[A]. Nucleotide acid sequence and deduced amino acid sequence. of the amorphadiene synthase encoding gene, between start and stop codon, (flanked with a *Nco* I and *Bam*H I site) obtained by PCR with the primers C and D.

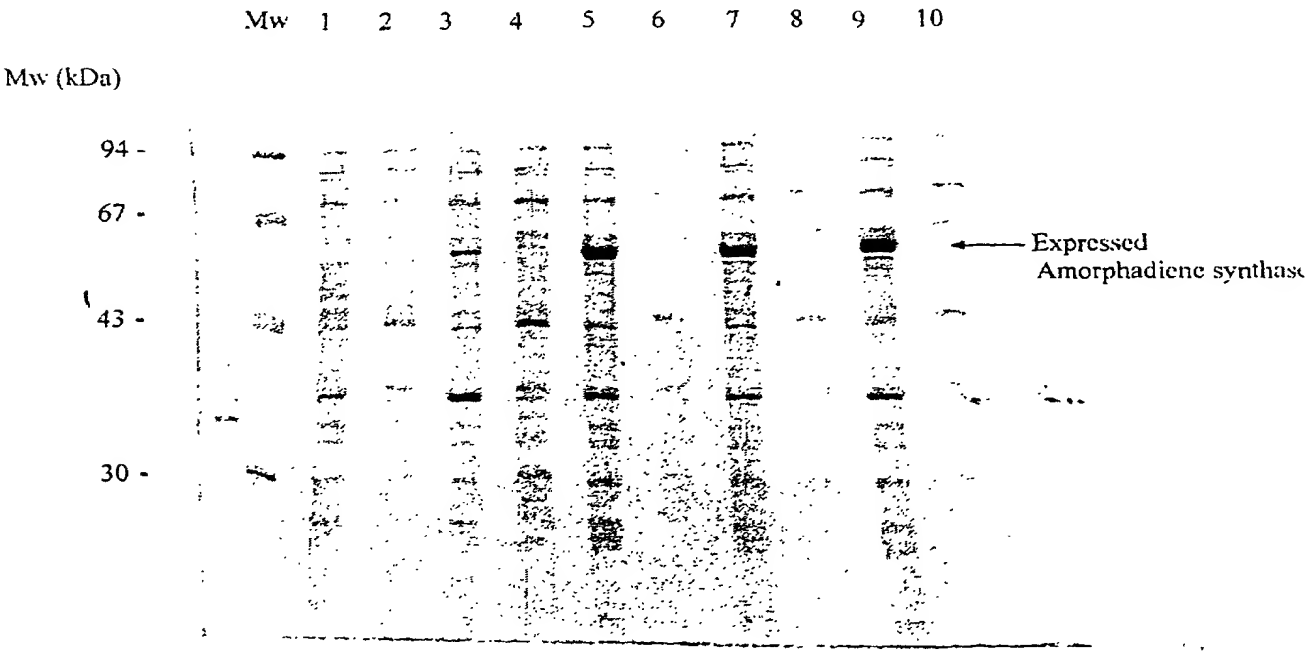



Fig. 10. SDS-PAGE gel: Mw = low Molecular Weight marker (Pharmacia Biotech), lane 1 and 2 resp. pellet and supernatant pET 11d (negative controle), lane 3 and 4 resp. pellet and supernatant TEAS in pET 11d (positive controle), lane 5,7,9 and 6,8,10 resp. pellet and supernatant amorphadiene synthase in pET 11d. All constructs were expressed in *E. coli* BL21(DE3).The lanes with the pellet fractions of TEAS in pET 11d (positive controle) and amorphadiene synthase in pET 11d showed a clear spot which was not present in the negative controle pET 11d.

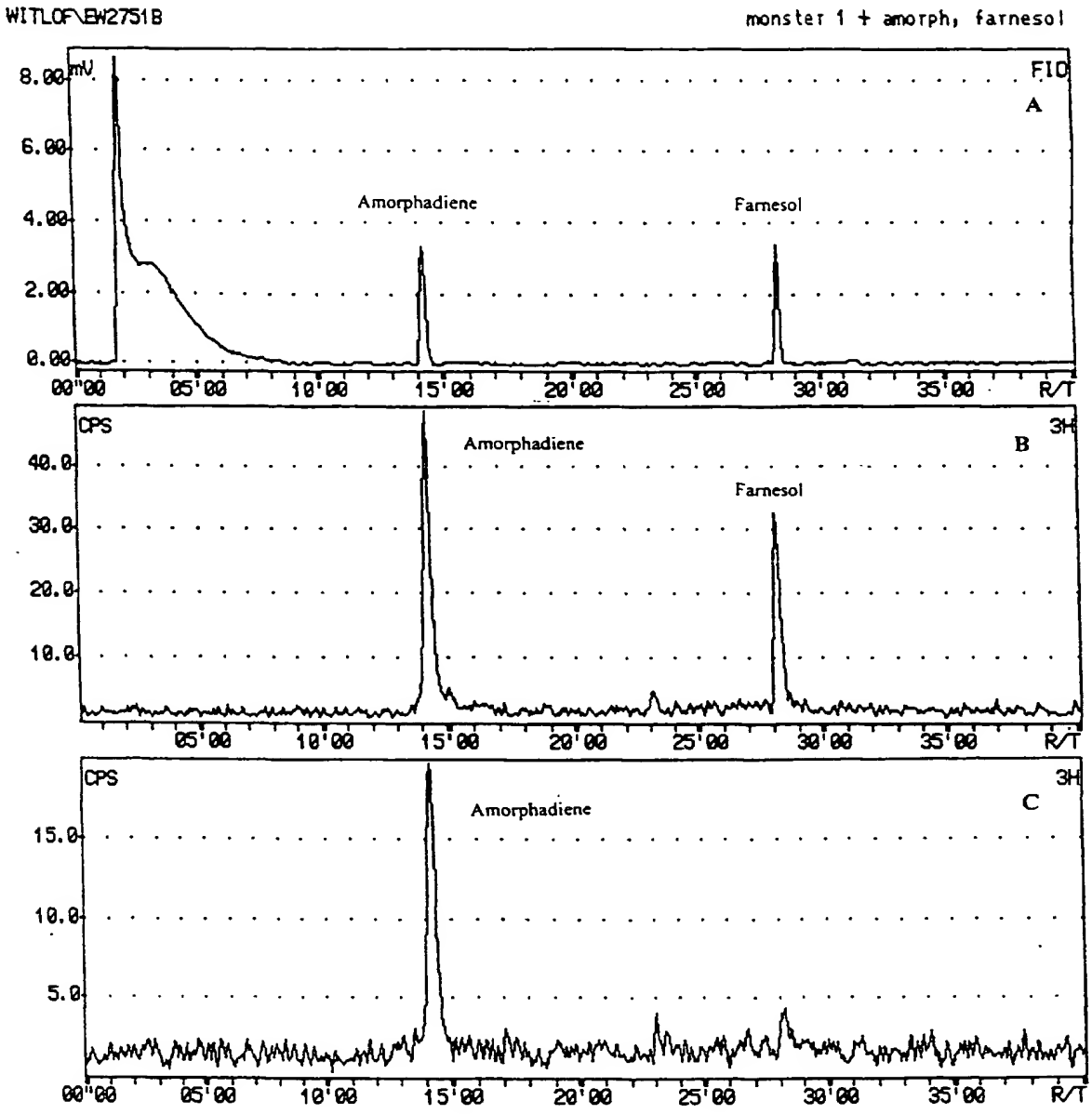



Fig. 11. A: Flame ionization detector (FID) signals of amorpha-4,11-diene and farnesol (references). Radio-GC chromatograms of the [$^3$H]-FPP-assays with. B: intact BL21(DE3) cells and C: the supernatant of sonicated BL21(DE3) cells, both transformed with the amorphadiene synthase encoding gene in the expression vector pET11d.

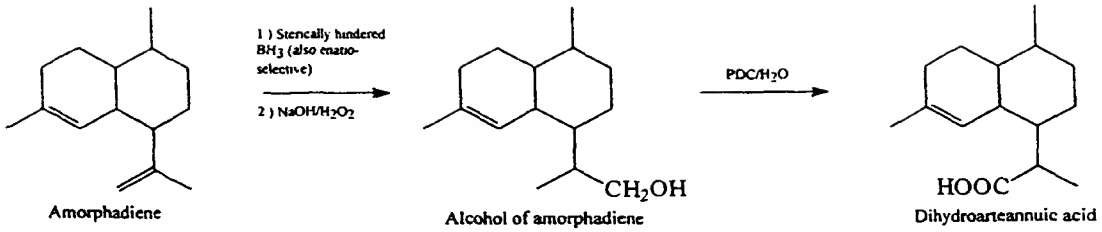



Fig. 12 Hypothetical chemical synthesis of dihydroaeteannuic acid using amorpha-4,11-diene as a precursor. The reaction consist of an enantio-, stereo- and regio selective (anti-markownikoff) hydroboration of amorphadiene with $BH_3$ followed by an oxydation of the formed trialkylboranes with $NaOH/H_2O_2$ yielding the alcohol. A mild oxidation of the alcohol to the acid can be obtained with **PCD** (pyridinium dichromate) without attacking the second bouble bond.

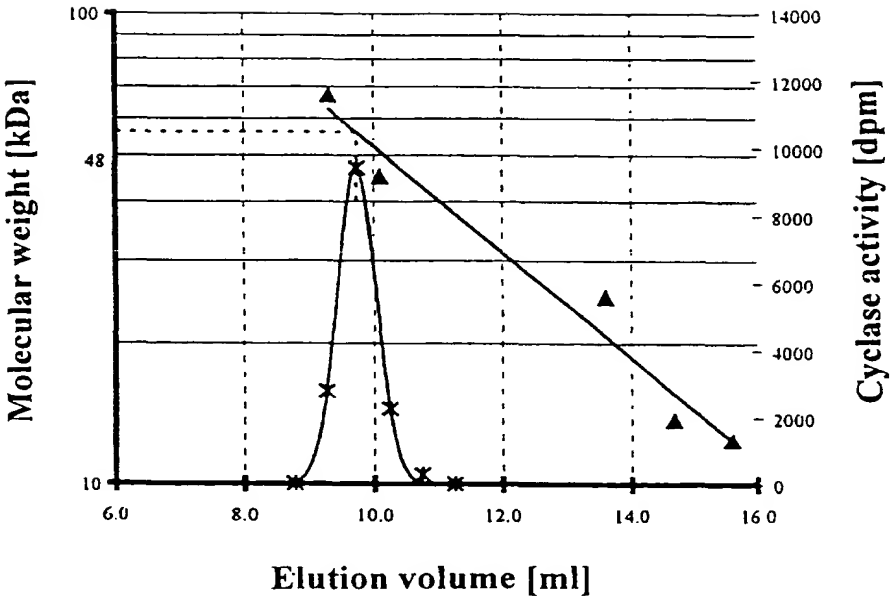



Fig. 13 Determination of the molecular weight of amorpha-4,11-diene synthase by size-exclusion chromatography (gel filtration). (-*-): activity curve; (-▲-): molecule weight markers: (—): molecular weight calibration line.